An improved 3D SIFT applied to estimate volume displacement field

Chengsheng Li[1,2]

[1] Department of Civil Engineering and Smart Cities, Shantou University, Shantou, 515000, China

[2] State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, Hubei 430000, China

E-mail: chengshengli@stu.edu.cn

Abstract

DVC method has great potential in calculating 3D displacement field and strain field, but it relies heavily on the initial value of displacement field, especially for complex and large deformation cases. 3D SIFT algorithm is a very effective method to obtain feature point matching results. However, although the current 3D SIFT algorithm can extract a large number of feature points, the number of points that can be matched is not enough. This is because the orientation of the key points is not properly calculated. Therefore, we use the matrix singular value decomposition (SVD) method instead of the matrix eigenvalue decomposition method to deal with the image structure tensor of the key point and use Gaussian smoothing to deal with the amplitude of the gradient to obtain a more reasonable and accurate key point direction. It was verified by CT triaxial test of granite residual soil. The results show that the improved 3D SIFT method increases the matching rate of feature points by 110%~180%, and the matching results provide more accurate and dense initial displacement field estimation results for the digital volume correlation (DVC) method.

Keywords: 3D SIFT, Image matching, SVD, CT, DVC

1. Introduction

Digital volume image correlation method (DVC) is widely used in the measurement of three-dimensional displacement field and strain field of materials [1-4]. It plays an important role in the three-dimensional view of strain localization of materials [5]. Whether local DVC methods or global DVC method, have high requirements on the initial value of the displacement field [6], especially for complex and large deformation cases, otherwise there may be decorrelation or non-convergence [7]. However, the DVC method relies heavily on the initial value of the three-dimensional displacement field, especially for complex and large deformation cases. Among various feature point matching methods, the Scale-invariant feature transform (SIFT) algorithm has significant effects under noise, rotation,

scaling, and multi-scale conditions[8]. Many scholars have conducted improvement research on SIFT, which has been widely used in many fields. However, the classical SIFT algorithm is suitable for processing two-dimensional images and cannot be directly applied to volume images.

In order to enable SIFT algorithm to effectively process volume images (such as CT and MRI imaging), many scholars have done a lot of work [10-14]. For example, Allaire, Stephane et.al[15] established the Hessian matrix in the key point region to describe the principal curvature, which is used to filter the extreme points belonging to the plane or cylinder. For three-dimensional volume images, since the key point orientation calculation method in the two-dimensional SIFT method cannot be directly used. Blaine Rister et al.[16] proposed to use the structure tensor of images to describe them, use matrix eigenvalue decomposition to calculate the key point orientation. When establishing local descriptors of key points, histogram statistical enrichment used by the 2D SIFT algorithm cannot be directly applied to 3D. Klaser, Alexander, et al. [17] try to use polyhedron statistical gradient vector distribution. On that basis, Blaine Rister et al.[16] put forward a positive 20-hadron statistical method to establish local descriptors of key points. The adopted regular 20-hedron can just divide the three-dimensional space equally, and the regular 20-hedron is the regular polyhedron with the largest number of faces. The improved 3D SIFT has good feature point matching capability.

Accurate key direction is very important in the 2D SIFT algorithm for matching feature points. The orientation of the key points is determined by the gradient's histogram statistics. To obtain more stable and effective matching results, the distribution of the histogram is also calculated by curve interpolation, and even several extreme value directions need to be calculated. The current 3D SIFT algorithm can calculate a large number of feature points, but due to the lack of key point direction calculation, the matching success rate is low, and a large number of key points are "wasted". In addition, high-density feature point matching results are crucial for 3D displacement field estimation of volume images, especially for the DVC algorithm, so improved 3D SIFT has broad application prospects.

2. 3D SIFT algorithm

2.1 Keypoint locations

Similar to the two-dimensional SIFT method, we obtain the Laplacian of Gaussian function:

$$l(\mathbf{x},\sigma) = I(\mathbf{x}) * \Delta g_{\sigma}(\mathbf{x}) \tag{1}$$

where σ is a Gaussian scaling parameter, $\mathbf{x}$ is the image space coordinate, $g_{\sigma}(\mathbf{x})$ is a Gaussian function.

This is approximated by convolution with the difference of Gaussians (DoG) function:

$$d(\mathbf{x},\sigma) = I(\mathbf{x}) * (g_{\sigma+\delta}(\mathbf{x}) - g_{\sigma}(\mathbf{x})) \tag{2}$$

where δ is a small constant.

In order to screen the local extremum more effectively, the maximum extremum of the image is taken as the standard, and the extremum exceeding α ratio is selected.

$$|d(\mathbf{x},\sigma)| > \alpha \max_{\mathbf{x},\sigma} |d(\mathbf{x},\sigma)| \tag{3}$$

In addition, to avoid the local extreme point being a plane or cylinder, the following judgment formula is adopted based on establishing the key point Hessian matrix[7]:

$$\frac{\mathrm{tr}(\mathbf{H})^3}{\det(\mathbf{H})} < \frac{(2t_{\max}+1)^3}{(t_{\max})^2} \tag{4}$$

where $t_{\max}$ is the principal curvatures.

2.2 Local orientations and corner detection

In the 2D SIFT algorithm, the dominant direction in the local gradient histogram is used as the orientation of the key points, but the direct expansion of this method to three dimensions cannot make the key points have rotation invariance. Therefore, Blaine Rister et al.[8] proposed to use the structure tensor of images to describe:

$$K = \int w(\mathbf{x})\nabla I(\mathbf{x})(\nabla I(\mathbf{x}))^T \, d\mathbf{x} \tag{4}$$

where $\nabla I(\mathbf{x})$ is the gradient of image I at location $\mathbf{x}$, approximated by finite differences, and $w(\mathbf{x})$ is a Gaussian window centered at the keypoint, the diameter of which is a constant multiple of the keypoint scale.

Using matrix eigenvalue decomposition:

$$K = R\Lambda R^T \tag{5}$$

We can get three eigenvectors R (r_1, r_2, r_3).

Because the vector distribution in the spherical region of the key points does not always conform to the ellipsoidal distribution, the feature vector obtained from the structure tensor can not describe the feature direction of the key points well, and a large number of key points can not be successfully matched.

To address this issue, we have made the following improvements:

(a) Gaussian smoothing of vector amplitude

To avoid the noise effect of the key local gradient vector, we use Gaussian smoothing, the value of σ is generally 0.5.

$$|\nabla I(\mathbf{x})|' = g_\sigma(|\nabla I(\mathbf{x})|) \tag{6}$$

It should be noted that Gaussian smoothing is not used when building local descriptors.

(b) Singular value decomposition (SVD) of structure tensors

When using 3D SIFT, we found that the orientation of the key points and the statistical principal axis of the gradient in the local descriptor need not be equal or parallel, but the angle between them must be fixed or conform to the same law. There is another kind of matrix decomposition called matrix singular value decomposition, which obtains a series of singular values. In most cases, the value of the singular value decays rapidly. For the structure tensor of the three-dimensional image, the singular value obtained by using matrix singular value decomposition has significant "criticality", and the first singular value accounts for a large "component". This "prominence" property far exceeds the characteristic of matrix eigenvalue decomposition, and this property is very much in line with the requirement of determining the key direction in the key points. Therefore, the base vector is obtained using a more stable Singular Value Decomposition (SVD) :

$$K = U \Lambda V^T \tag{7}$$

We get the "left" basis vector U and the "right" basis vector V.

In general, we directly use the basis vector U (u_1, u_2, u_3) as the three key directions of the key points. If the basis vectors are not orthogonal, further orthogonalization is required.

2.3 Gradient histograms

The local descriptor gradient statistical method used in 2D SIFT cannot be directly extended to 3-dimensional space. As shown in Fig.1, the positive 20-hedron composed of 12 vertices is the positive polyhedron with the largest number of faces. It can just divide the three-dimensional space equally, so it can be used to count local descriptors [16]. The Möller-Trumbore algorithm [18] is used to determine whether the vector intersects the specified triangular plane. If it intersects, the weight accumulation of the vector in the direction of each vertex of the triangular surface is determined according to the area ratio between the position of the intersection point and the vertex of each triangular surface.

When a local descriptor of a key point is established, a spherical region of radius 4σ is taken as the center of the key point and divided into subsets 4×4×4. Then let the initial coordinate system rotate to the key point orientation calculated by 2.2. Finally, the vector distribution in the spherical region is calculated by using the above normal 20-hedral statistical method. A positive 20-hedron with 12 vertices, resulting in a local descriptor of $4^3 \times 12 = 768$ components in total.

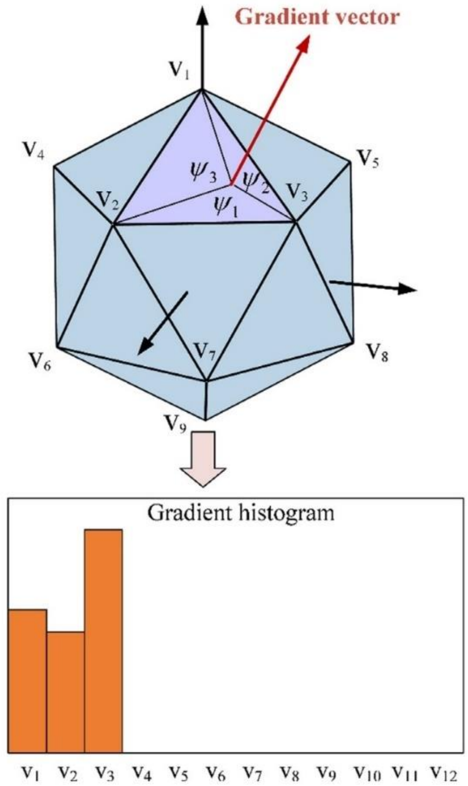

Fig. 1 Illustration of how a gradient vector of the voxel is assigned to three vertices of the congruent triangle intersected with it [6].

2.4 matching

In our matching, we go straight to brute force. Only keep matches in which the ratio of vector angles from the nearest to the second nearest neighbor is less than ratio η. In our study, we chose η=0.65 as the screening threshold. When the descriptors between two key points meet this condition, it means that the two key points are an optimal matching point, and the difference in coordinate values between them is also the displacement value between them. The displacement values of some columns of key points can be combined into a three-dimensional displacement field, which can be used as the initial displacement field of the DVC algorithm.

3 Experiment

We chose the natural granite residual soil shown in Fig. 2 for a CT triaxial test to validate the effectiveness of the proposed method. Prior to the CT triaxial test, a cylindrical sample of natural granite residual soil with a diameter of 60 mm and a height of 120 mm was prepared and positioned in the pressure chamber of the triaxial compression tester. The confining pressure was set at 50 kPa, and the drainage condition was established as undrained. The samples at different strain levels were scanned via CT to generate a sequence of volume images corresponding to axial strains of 4%, 6%, 9%,

and 11.5%. The size of the volume images was 2000 × 2000 × 3000 voxels. For calculation, the volume image is compressed to 300 × 300 × 458 voxels.

The calculation parameters of 3D SIFT are as follows: Octaves = 3, Scales = 5, σ = 1.3, α = 0.1, t_{max} = 20, η = 0.65.

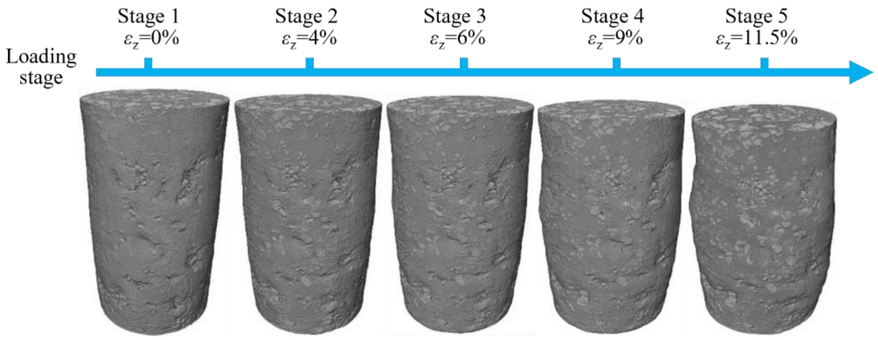

Fig. 2 CT reconstruction results of granite residual soil.

3.1 Rotational experimental

As shown in Fig. 3, CT reconstruction results of the first stage were taken as an example. The Z-axis was taken as the rotation axis, rotated 5°, 15°, 30°, and 45°, and then 3D SIFT was used to calculate feature points, and finally, feature points under different rotation angles were matched, as shown in Table 1.

(1) The feature points obtained by the improved 3D SIFT are the same, but the matching rate is increased by nearly 110%~180%. Compared with the eigendirections obtained by matrix eigenvalue decomposition, the rotation invariance of the directions obtained by SVD is more stable.

(2) All matching rates decay with the increase of rotation Angle, but the attenuation of the improved 3D SIFT algorithm is smaller. This phenomenon further indicates that the rotational invariance of the direction obtained by SVD is more stable.

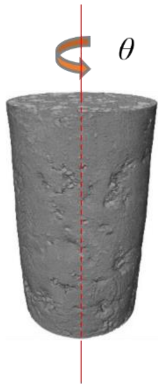

Fig. 3 Volume images reconstructed by CT are rotated at different angles along the Z-axis.

Table 1 Rotate key points and match results in simulation data.

Rotation Angle	Keypoints	3D SIFT [16]		Improved 3D SIFT		Increased
		Match points	Match ratio	Match points	Match ratio	
0°	94586	-	-	-	-	-
5°	94392	26048	27.54%	54837	57.76%	110.53%
15°	97885	16924	17.89%	42762	45.04%	152.67%
30°	95923	12922	13.66%	35159	37.03%	172.09%
45°	95044	10079	10.66%	28461	29.98%	182.38%

3.2 Displacement field estimation

Taking the CT reconstruction results of the five stages in Fig 1 as an example, the 3D SIFT algorithm is used to calculate the feature points and matching results in 1→N mode, as shown in Table 2, Fig 4, Fig 5, and Fig 6.

(1) The improved 3D SIFT algorithm increases the matching rate of feature points by 130%~140%, which can more accurately predict the local displacement field of the sample, especially for large deformations such as particle rotation and slip.

(2) When the axial strain is small, there is a small deformation between the key points calculated by the 3D SIFT, including stretching, compression, and rotation, which results in a high matching rate of the key points between the different loading stages. However, because the key directions obtained by SVD are more stable, the final key point matching is higher than that of traditional 3D SIFT methods.

(3) With the increase of axial strain, shear band failure occurs in the middle section of the sample, and the matching points calculated by the two 3D SIFT methods are too few. This is because there may be severe deformation or even rotation of particles in the shear zone region, resulting in drastic changes in the texture features reconstructed by CT scanning, resulting in difficulty in successfully matching local descriptors of key points. However, the matching points obtained by the improved 3D SIFT can still be more than doubled, that is, it can provide more accurate displacement field estimation results.

Table 2 Key points and matching results of different loading stages.

Loading stage	Keypoints	3D SIFT[8]		Improved 3D SIFT		Increased
		Match points	Match ratio	Match points	Match ratio	
1	94586	-	-	-	-	-
2	99768	17717	18.73%	40919	43.10%	130.96%
3	93989	15043	15.90%	34750	36.60%	131.00%
4	95220	9827	10.39%	24152	25.44%	145.77%
5	80656	4433	4.68%	10549	11.11%	137.97%

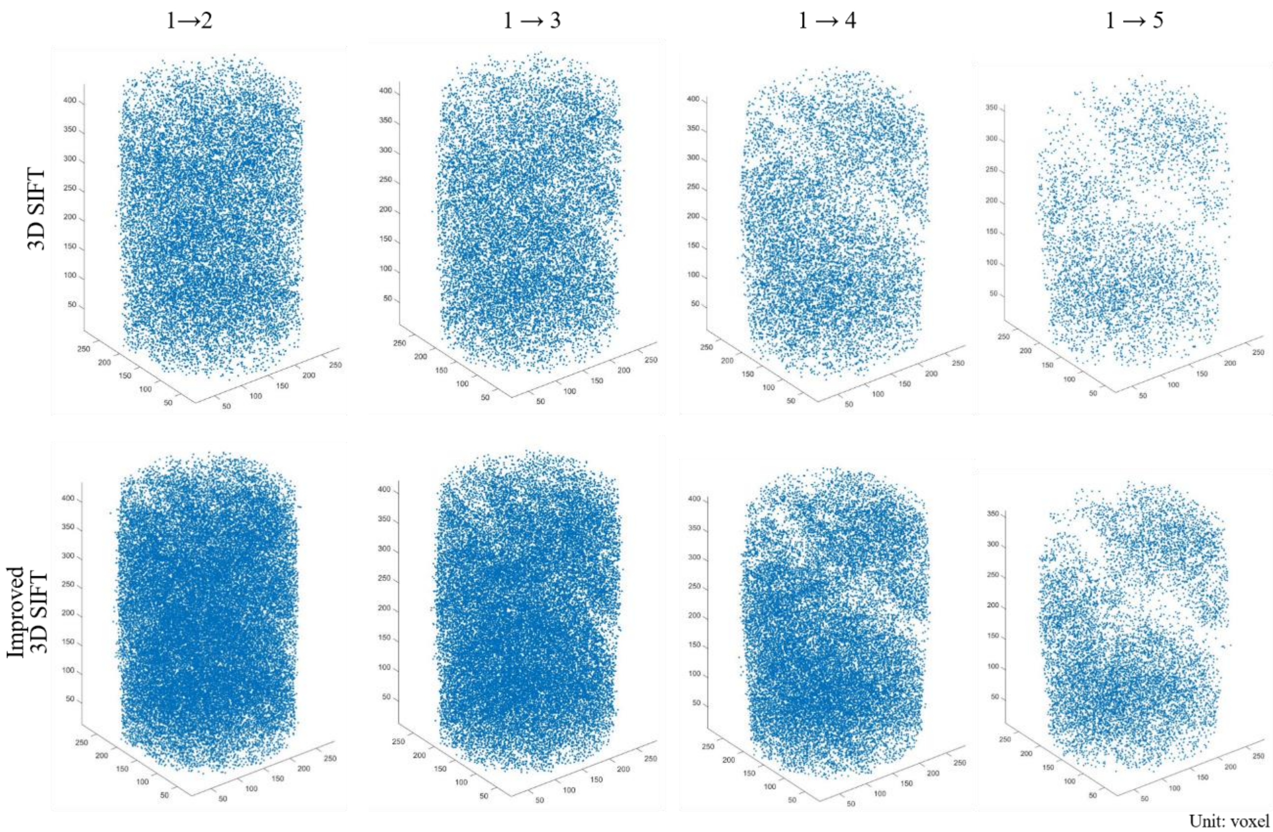

Fig. 4 Three-dimensional spatial distribution of matching results of feature points.

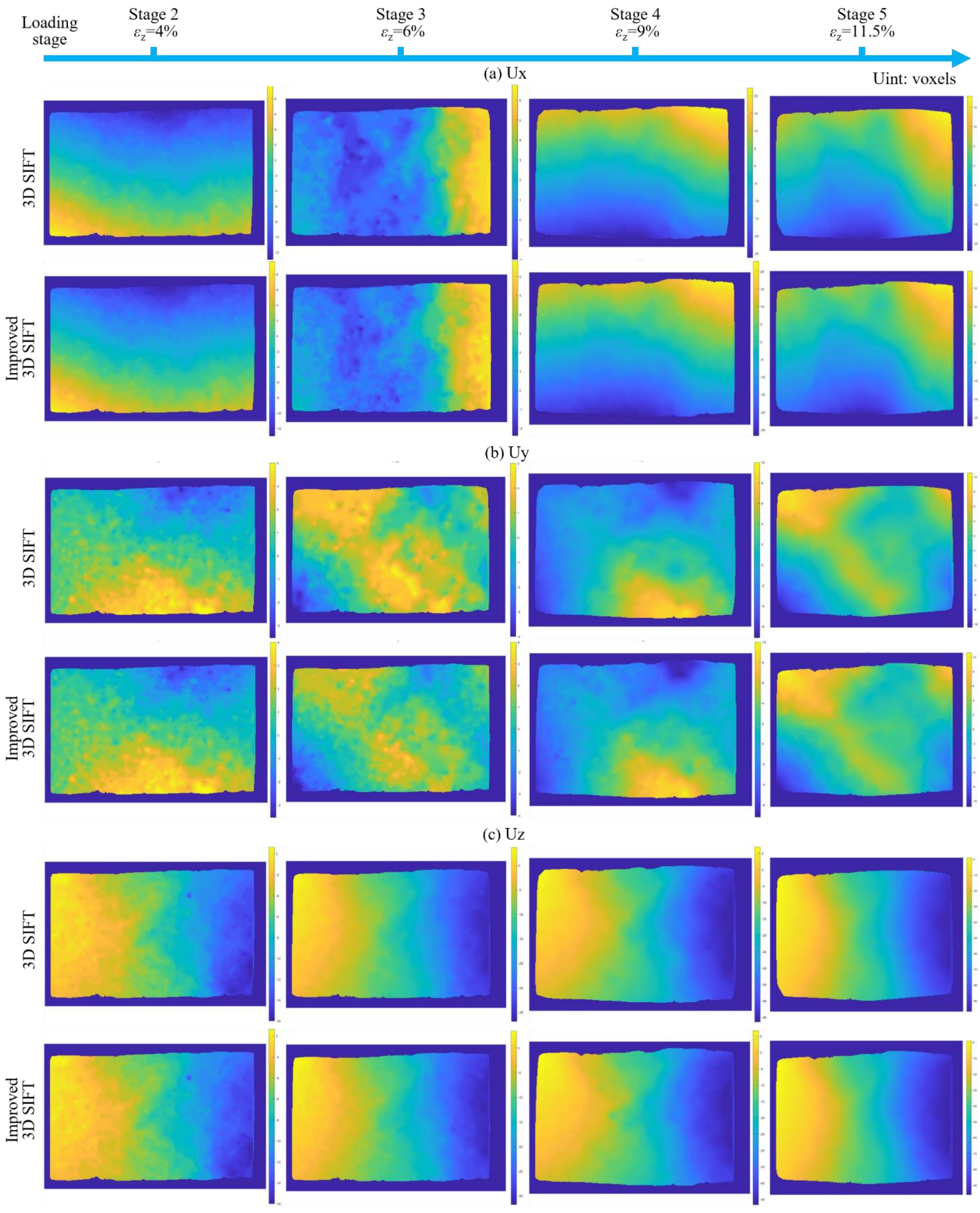

Fig. 5 Displacement cloud image with x = 150 voxels

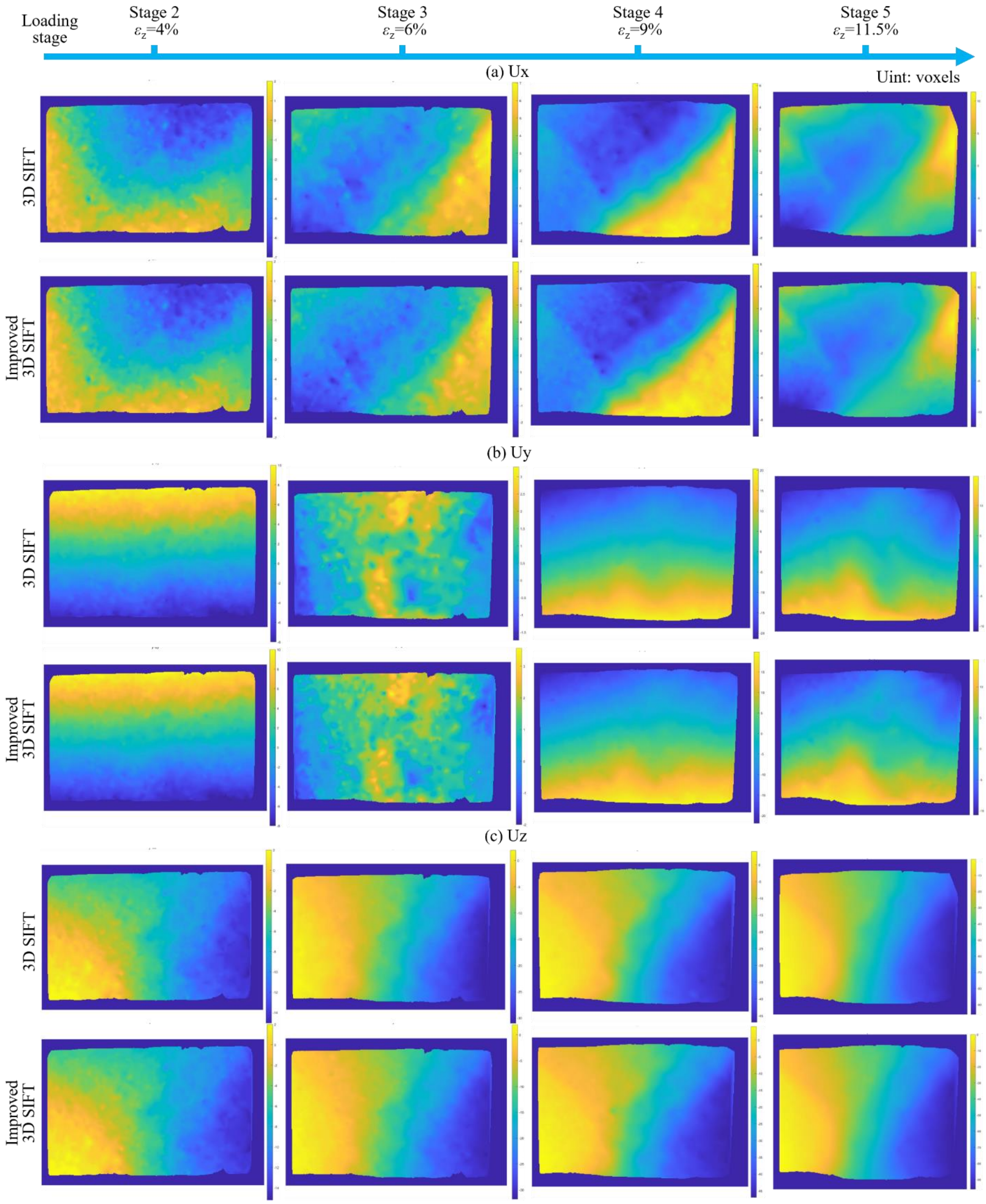

Fig. 6 Displacement cloud image with y = 150 voxels

4 Conclusion

According to the above experimental analysis, we preliminarily verify that the basis vector obtained by

processing the image structure tensor with matrix eigenvalue decomposition is not good enough, and its rotation invariance is not good enough in the rotation numerical analysis. The basis vector obtained by SVD decomposition has a more robust rotation invariance, so it can improve the matching rate of river key points in different rotation tests by 110%~180%, and make full use of the calculated key points. The improved 3D SIFT is very suitable for the initial displacement field estimation in the DVC algorithm.

On the other hand, we only used one kind of CT image for numerical verification and tested the effectiveness of the improved 3D SIFT algorithm under different materials and different scanning accuracy, which needs further in-depth research. In particular, how to assess the accuracy of the direction of key points.

Acknowledgments

This work was supported by the Building Fund for the Academic Innovation Team of Shantou University (CN) (NTF21017), the Special Fund for Science and Technology of Guangdong Province in 2021 (STKJ2021181).

References

[1] Wang B, Pan B. Anisotropic self-adaptive digital volume correlation with optimal cuboid subvolumes. Meas Sci Technol. 2019;30:115008.

[2] Yang J, Hazlett L, Landauer AK, Franck C. Augmented Lagrangian Digital Volume Correlation (ALDVC). Experimental Mechanics. 2020;60:1205-1223.

[3] Bi Y, Liu H, Mao L, Liu J, Liu Y, Zuo J, et al. Investigation into debonding of single polypropylene fiber pullout in concrete using X-ray microtomography and mechanically regularized digital volume correlation. Materials and Structures. 2024;57:188.

[4] Wu M, Wang J, Pan B, Yin Z-Y. Particle tracking–aided digital volume correlation for clay-sand soil mixtures. Géotechnique. 2024:1-32.

[5] Li C, Kong L, Zhang B, Liu Z, Shu R, Li T. Determination of damage evolution characteristics in granite raesidual soil shear bands by micro-CT-based advanced digital volume correlation. Engineering Geology. 2024;333:107505.

[6] Yang J, Huang J, Jiang Z, Dong S, Tang L, Liu Y, et al. 3D SIFT aided path independent digital volume correlation and its GPU acceleration. Optics and Lasers in Engineering. 2021;136:106323.

[7] Wang B, Pan B. Incremental digital volume correlation method with nearest subvolume offset: An accurate and simple approach for large deformation measurement. Advances in Engineering Software. 2018;116:80-88.

[8] Lowe DG. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision.

2004;60:91-110.

[9] Zhou G, Li H, Gao E. Overview on Interest Point Detection Over 40 Year's Development: A Review. IEEE Sensors Journal. 2023;23:26942-2651.

[10] Cheung W, Hamarneh G. N-SIFT: N-dimensional scale invariant feature transform for matching medical images. IEEE International Symposium on Biomedical Imaging: from Nano to Macro 2007.

[11] Scovanner P, Ali S, Shah M. A 3-dimensional sift descriptor and its application to action recognition. Acm International Conference on Multimedia 2007.

[12] Ni D, Qu Y, Yang X, Chui YP, Wong TT, Ho SSM, et al. Volumetric Ultrasound Panorama Based on 3D SIFT. 2008.

[13] Wu H, Wang D, Shi L, Wen Z, Ming Z. Midsagittal plane extraction from brain images based on 3D SIFT. Physics in Medicine and Biology. 2014;59:1367-1387.

[14] Xu Y, Xu C, Kuang X, Wang H, Chang EI-C, Huang W, et al. 3D-SIFT-Flow for atlas-based CT liver image segmentation. Medical Physics. 2016;43:2229.

[15] Allaire S, Kim JJ, Breen SL, Jaffray DA. Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis. IEEE Computer Society Conference on Computer Vision & Pattern Recognition Workshops 2008.

[16] Rister B, Horowitz MA, Rubin DL. Volumetric Image Registration From Invariant Keypoints. IEEE Transactions on Image Processing. 2017;26:4900-4910.

[17] Kläser A, Marszalek M, Schmid C. A Spatio-Temporal Descriptor Based on 3D-Gradients 2008.

[18] Möller T, Trumbore B. Fast, Minimum Storage Ray-Triangle Intersection. Journal of Graphics Tools. 2005;2.